A posterior probability base sequential binormal test method for verification of success ratio P

Liu, Qi; An, Min

DOI:
10.1260/1708-5284.8.4.313

A posterior probability based sequential binomial test method for verification of success ratio p

by

Q. Liu and M. An

reprinted from

WORLD JOURNAL OF ENGINEERING

VOLUME 8 NUMBER 4 2011

MULTI-SCIENCE PUBLISHING COMPANY LTD.

World Journal of Engineering 8(4) (2011) 313-324

World Journal of
Engineering

A posterior probability based sequential binomial test method for verification of success ratio *p*

Q. Liu[1] and M. An[2,*]

[1]*School of Information Systems and Management, National University of Defence Technology, Changsha, P. R. China, 410073*
[2]*School of Civil Engineering, University of Birmingham, Birmingham B15 2TT, UK*
**E-mail: M.An@bham.ac.uk*

(Received 2 August 2011; accepted 8 September 2011)

Abstract

Testing the lifetimes of components or products by means of the verification of probability of success ratio p on which to base a statistical probability of the binomial distribution is often a costly and difficult undertaking. Sometimes tests cannot reach at a desirable target, particularly, in the reliability assurance tests. A Bayesian sequential binomial test model (BSBTM) is proposed for obtaining the composite hypothesis of p, which posterior criteria are taken into consideration. In order to get robust decision criteria and closed continuation-sampling regions, a modified Bayesian sequential binomial test model (MBSBTM) is also developed. By using BSBTM and MBSBTM, the upper and lower boundaries of a continuation-sampling region can be determined and the decision criteria can be made. A simulation method of calculating the average sample number (ASN) by using MBSBTM is also presented in this paper. Two case examples are used to demonstrate the proposed BSBTM and MBSBTM methodologies. The results indicate that by using MBSBTM sample numbers can be decided effectively and efficiently in the reliability tests.

Key words: *Bayesian, BSBTM, MBSBTM, Sequential test, Posterior probability*

1. Introduction

The purpose of the verification of probability of success p in the lifetime tests, especially in the assurance tests of binomial distributed components or products, is to determine the sample numbers to satisfy the consumer's requirements of the quality of products before assembly and manufacture. Many of sampling test methods, such as fixed-sample tests (Desu and Raghavarao, 1990; Hines and Montgomery, 2003) and sequential probability ratio tests (SPRT) (Ghosh, 1970; Kapur and Lamberson, 1977; Wetherill and Glazebrook, 1986) are comparatively mature methods and usually depend on a hypothesis testing of p. However, in many circumstances, the application of these methods may not give satisfactory results in determining average sample number (ASN) due to the unknown parameters in the prior information before actual tests. Therefore, it is essential to develop new methods to predict the numbers of samples and to verify the probability of success p in an acceptable way under various environments

ISSN:1708-5284

where such mature methods cannot be effectively or efficiently applied. Bayesian method has been proved to be an effective and efficient method in reliability assurance tests (Hamada *et al*., 2008; Koch, 2007). Incorporating Bayesian method into SPRT method (Guo *et al*., 2008) proposed a method of sequential posterior odd test (SPOT) for sequential tests, which acquires not only reliable results, but also makes fully use of prior information in order to reduce the required ASN. By comparing fixed-sample test method (Desu and Raghavarao, 1990; Hines and Montgomery, 2003) with SPOT (Guo *et al*., 2008; Ghosh, 1970), the ASN of SPOT in a sequential test is smaller than a fixed-sample test for the given risks α and β of the producer and consumer, respectively. α and β are defined as (Lefebvre, 2006; Dekking, 2005)

$$\begin{aligned}\alpha &= P(\text{Error of Type I}) \\ &= P(\text{Reject } H_0 | H_0 \text{ is true})\end{aligned} \tag{1}$$

$$\begin{aligned}\beta &= P(\text{Error of Type II}) \\ &= P(\text{Accept } H_0 | H_0 \text{ is false})\end{aligned} \tag{2}$$

where H_0 is null hypothesis of parametric test. A Type I error occurs when H_0 is rejected due to an incorrect decision made by the producer, while Type II error occurs when H_0 is not rejected due to an incorrect decision made by the consumer.

However, because a Bayesian assurance test is a sequential test, a decision usually needs to be made at a time on whether to continue testing or stop testing when a sampling has been completed. Current methods only address the determination of the posterior risks α and β, but these methods do not provide clear decision criteria based on actual testing conditions and situations. For example, Jiang *et al.,* (2009) introduced a Bayesian SPRT method in the design and planning of tests. On the basis of the posterior risks α and β it just simply states that if $P(H_0 \mid x)$ is far larger than $P(H_1 \mid x)$, then accept H_0, and if $P(H_0 \mid x)$ is far smaller than $P(H_1 \mid x)$, then accept H_1. By using a lognormal distribution, Guo *et al.,* (2008) proposed a censored SPOT method for the verification of the mean time to repair (MTTR). In this method, for a simple hypothesis (*i.e.* $H_0: p = p_0$ vs H_1: $p = p_1$), it is easy to calculate the posterior risks and design a sequential sampling plan, as well as to determine the acceptance region, rejection region and continuation region. But for a composite hypothesis (*i.e.* H_0: $p \geq p_0$ vs. $H_1: p < p_1$, where $p_1 \leq p_0$), this method may not give a satisfactory sequential sampling plan. Furthermore, for a sequential sampling test with binomial distributions, the fixed-sample test method introduced by Hamada *et al.,* (2008) can only make a test plan for given posterior risks α and β. However, for composite hypothesis of failure rates (*i.e.* $H_0: \theta \leq \theta_0$, where θ is failure rate and θ_0 is a critical value for the failure rate that the producer claims to meet to satisfy the consumer's requirements of quality of products), Barnett (Barnett, 1972; Hoff, 2009) proposed a Bayesian sequential life test model that considers posterior probability, but the engineering requirement of the verification of success ratio does not take into consideration. Therefore, the existing methods have at least the following problems.

- The existing models only take the posterior risks of Types I and II errors (Ghosh, 1970; Kapur and Lamberson, 1977; Koch, 2007) into account, but do not consider the posterior probability of H_0 or H_1. Therefore, the results of hypothesis test may not satisfy the 'small probability event principle' (Li and Li, 2008).
- For a composite hypothesis of success ratio, p, the existing models cannot be applied to make an effective test plan before testing (Li and Li, 2008).

This paper presents a Bayesian sequential binomial test model (BSBTM) and a modified Bayesian sequential binomial test model (MBSBTM), which the posterior probability with composite hypothesis of p is taken into consideration. By using BSBTM and MBSBTM, the upper and lower boundaries of continuation-sampling region can be determined so that the sequential sampling plan and decision criteria can be made effectively and efficiently. Two case examples are used to demonstrate the applications of BSBTM and MBSBTM.

2. The assumptions of a sequential test

2.1. The assumptions of parametric tests

In the reliability assurance tests of binomial distributed components or products, the consumer is interested in whether or not the success ratio p satisfies a given requirement p_0, which is equivalent to test the following statistical hypothesis

$$H_0: p \geq p_0 \text{ and } : H_1 : p < p_0 \tag{3}$$

Supposing the prior distribution of p is $\pi(p)$. If it is a conjugate distribution (prior distributions that take the same functional form as the posterior distribution are called conjugate prior distributions (Hamada *et al.*, 2008), the prior distribution will be a Beta distribution $B(a, b)$, where a and b are hyper-parameters, *i.e.* a is the pseudo-success-number and b is the pseudo-failure-number in the prior reliability test. The probability density function (PDF) of $B(a, b)$ is defined as

$$f(p \mid a,b) = \frac{p^{a-1}(1-p)^{b-1}}{\int_0^1 p^{a-1}(1-p)^{b-1}\,dp} \tag{4}$$

Let n denote the actual sample number and s denote the successful sample number in a sequential test, where $n = 0, 1, 2, \cdots$ and $s = 0, 1, 2, \cdots, n$. The posterior distribution of p will be $\pi(p \mid (n, s))$, which can be obtained by Bayesian formula (5) (Desu and Raghavarao, 1990)

$$\pi(p \mid (n,s)) = \frac{\pi(p)L(p \mid (n,s))}{\int_0^1 \pi(p)L(p \mid (n,s))dp} \tag{5}$$

where $L(p \mid (n,s)) = p^s(1-p)^{n-s}$ is likelihood function.

Particularly, when the prior distribution is a Beta distribution, $B(a, b)$, the posterior distribution will be $B(a + s, b + n - s)$ (Ghosh *et al.*, 2006; Hoff, 2009). For example, assume the prior distribution is $B(500, 50)$ and actual and successful sample numbers are (160, 150). The posterior distribution will be $B(650, 60)$. If actual and successful sample numbers are (160, 160), then posterior distribution will be $B(660, 60)$. The prior and posterior distributions are shown in Figure 1.

2.2. The process of a sequential sampling test

As described earlier in this paper, because the sample number and total operating time in the sequential sampling test cannot be pre-assigned, a decision of carrying on next testing or not should be made just after each sampling completed rather than at the end of all samples being tested in order to obtain reliable results and reducing testing time and resources (Kin *et al.*, 1997).

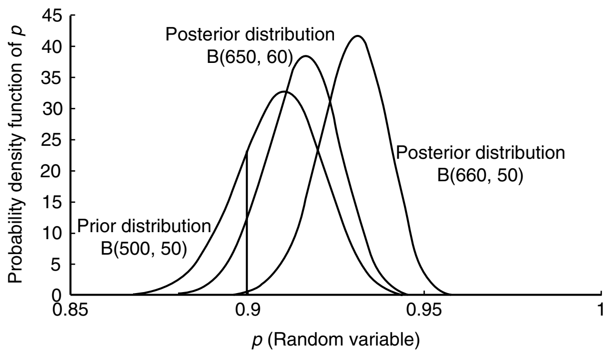

Fig. 1. Prior and posterior distributions.

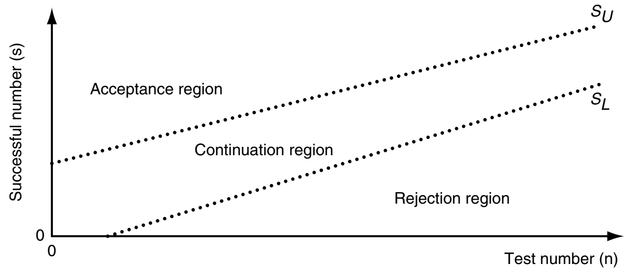

Fig. 2. Sequential sampling process of binomial distributed components or products.

Let n and s denote current actual and successful sample numbers, and s_U and s_L denote upper and lower boundary of continuation-sampling regions, respectively. The steps of a sequential sampling test for the hypothesis test of p are as follows:

Step 1: Divide the sample space into three mutually exclusive regions: rejection region, acceptance region and continuation region according to the given prior distribution.

Step 2: Calculate s_U and s_L of continuation region based on the given actual sample number n as shown in Figure 2.

Step 3: Verify the value of n, s_U and s_L.

- If $s_U \le s \le n$, then stop testing and accept H_0.
- If $0 \le s \le s_L$, then stop testing and reject H_0.
- If the testing result is 'success', then let $s = s + 1$ and , if $s_L < s < s_U$, then let $n = n+1$, select another sample to test (continue sampling), repeat *Step 3*.

For example, assume $n = 50$, $s_L = 34$ and $s_U = 45$. If $s = 40$, according to the above decision criteria, $34 < s < 45$, more tests need to be conducted.

3. Proposed Bayesian sequential binomial test model (BSBTM)

3.1. Fundamentals of BSBTM

On the basis of the assumptions described in section 2.1, when (n,s) have been gained, the posterior distribution $\pi(p \mid (n,s))$ can be obtained by using Bayesian formula (5) (Koch, 2007). According to reliability test hypothesis (3), the posterior probability of H_0 and H_1 can be calculated by

$$P(H_0 \mid (n,s)) = P(p \ge p_0 \mid (n,s)) = \int_{p_0}^{1} \pi(p \mid (n,s))\,dp \quad (6)$$

$$P(H_1 \mid (n,s)) = P(p < p_0 \mid (n,s)) = \int_{0}^{p_0} \pi(p \mid (n,s))\,dp \quad (7)$$

Considering the 'small event probability principle' (Hoff, 2009) in a hypothesis test, the decisions can be made as follows:

(1) If $P(H_0 \mid (n, s))$ is sufficiently small, for example $P(H_0 \mid (n, s)) < 0.005$, then H_0 has to be rejected. If H_0 is rejected, according to the definition of α (1), Type I error may occur. Therefore, the risk α of the producer can be defined as a sufficiently small number. In this case, the decision criterion will be: if $P(H_0 \mid (n, s)) < \alpha$ or $P(H_1 \mid (n, s)) \ge 1-\alpha$, then reject H_0, and the probability of Type I error is smaller than α as shown in Figure 3(a).

(2) If $P(H_1 \mid (n, s))$ is sufficiently small, for example $P(H_1 \mid (n, s)) < 0.05$, then H_1 has to be rejected, in other words, H_0 can be accepted. According to the definition of β (2), if H_0 is accepted, Type II error may occur. Therefore, the risk β of the consumer can be defined as a sufficiently small number. In this case, the decision criterion will be: if $P(H_1 \mid (n, s)) < \beta$ or $P(H_0 \mid (n, s)) \ge 1-\beta$, then accept H_0, and the probability of Type II error is smaller than β as shown in Figure 3(b).

In any cases, because $P(H_1 \mid (n, s))$ is a strictly monotone decreasing function of s while $P(H_0 \mid (n, s)) = 1-P(H_1 \mid (n, s))$ is a strictly monotone increasing function of s (see Appendix), if $\alpha + \beta < 1$, then the given (n, s) does not meet the above criteria 1) and 2). But, s in anyway satisfies either

$$\beta \le P(H_1 \mid (n,s)) < 1-\alpha \quad (8)$$

or

$$\alpha \le P(H_0 \mid (n,s)) < 1-\beta \quad (9)$$

Therefore, for any given (n, s) that does not meet the above criteria 1) and 2), the following criterion needs to be applied.

(3) If posterior probability of H_1 satisfies Inequality (8) or posterior probability of H_0 satisfies Inequality (9), then continue sampling.

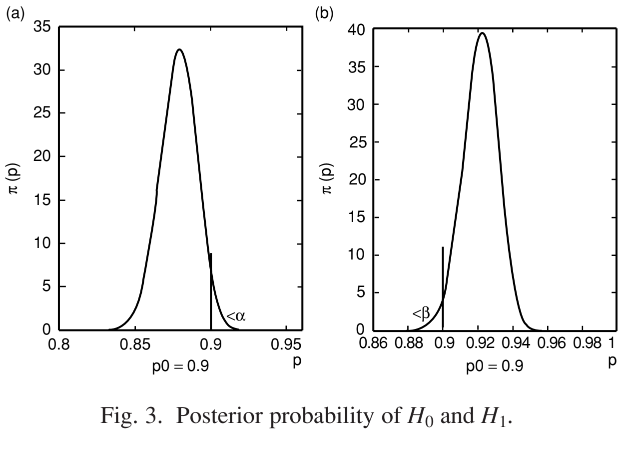

Fig. 3. Posterior probability of H_0 and H_1.

3.2. Solution of BSBTM

Supposing the prior distribution of p is the Beta distribution $B(a, b)$ and actual tests are (n, s), then the posterior distribution of p is $B(a+s, b+n-s)$. In this case, if H_0 is rejected, then the posterior probability of H_1 satisfies

$$P(H_1 \mid (n,s)) = \int_{0}^{p_0} f(p \mid a+s, b+n-s)\,dp \ge 1-\alpha \quad (10)$$

for a given n, $G(s)$ can be defined as

$$G(s) = \int_{0}^{p_0} f(p \mid a+s, b+n-s)\,dp \quad (11)$$

Obviously, for given a, b and, n, $G(s)$ is a strictly monotone decreasing function of s (see Appendix). In this case, the values of $G(s)$ will be more and more small with increasing s, the results will be either

(1) Let s_L denotes the maximum successful sample number s that satisfies Inequality (10), then s_L is the lower boundary of continuation region for a given n as shown in Figure 2.

(2) For a given n, any s ($s = 0, 1, 2, \cdots, n$) does not satisfy Inequality (10), let $s_L = -1$ by means of more tests needed for the verification of p. In this circumstance,

$$s_L = \max\{-1, \max\{s \mid G(s) \ge 1-\alpha\}, s = 1, 2, \cdots, n\} \quad (12)$$

However, if H_0 is accepted, then the posterior probability of H_1 must satisfy

$$P(H_1 \mid (n,s))$$
$$= \int_0^{p_0} f(p \mid a+s, b+n-s)dp < \beta \quad (13)$$

Similarly, for a given s (s = 0, 1, 2, ···, n), the values of $G(s)$ will be more and more small with increasing s, the results will be either

(1) Let s_U denotes the minimum successful sample number s that satisfies Inequality (13). Therefore, s_U is the upper boundary of continuation region for a given n as shown in Figure 2.

(2) For a given n, any s (s = 0, 1, 2 ···, n) does not satisfy Inequality (13), let $s_U = n + 1$ by means of more tests needed for the verification of p. In this circumstance,

$$s_U = \min\{n+1, \min\{s \mid G(s) < \beta\}, s = 1, 2, \cdots, n\} \quad (14)$$

Considering inequalities (10) and (13), for a given n, the following decision criteria can be obtained

- If $s_U \le s \le n$, then stop testing and accept H_0.
- If $0 \le s \le s_L$, then stop testing and reject H_0.
- If $s_L < s < s_U$, then select another sample and continue testing.

According to Formulas (12) and (14), obviously, when $\alpha + \beta < 1$, for any possible sample number n, $s_L \ne s_U$, the continuation region always exists as shown in Figure 2.

3.3. Censored BSBTM

As discussed in section 3.2, when $\alpha + \beta < 1$, for a given n, if $s_L < s < s_U$, then more sampling tests need to be conducted continuously. However, in many cases, there are only N censored samples in hand. Therefore, when $n = N$, the sequential test, however, must be terminated and a decision has to be made. In this case, the following decision criteria can be applied:

(1) Control β to protect the consumer's benefit. If $n \le N - 1$, the decision criteria are as same as described in section 3.2. If $n = N$, the decision criteria are
   - If $s_U \le s \le N$, stop testing and accept H_0.
   - If $s < s_U$, stop testing and reject H_0.

(2) Control α to protect the producer's benefit. If $n \le N - 1$, the decision criteria are as same as described in section 3.2. If $n = N$, the decision criteria are
   - If $0 \le s \le s_L$, stop testing and reject H_0.
   - If $s > s_L$, the stop testing and accept H_0.

(3) Compromise α and β. The decision criteria are: if $n \le N - 1$, the decision criteria are same as described in section 3.2; if $n = N$, let $s_M = (s_L + s_U)/2$, the decision criteria are
   - If $s \ge s_M$, stop testing and accept H_0.
   - If $s < s_M$, stop testing and reject H_0.

3.4. Numerical example 1

Considering the verification test of an engine starter, the prior distribution of p derived from previous test is B (500, 50) and $\alpha = \beta = 0.05$. Reliability requirement is $p_0 = 0.9$. Assume 1000 engine starters, i.e. $N = 1000$ available for reliability tests. For a given n (n = 1, 2, ..., N = 1000), s_L and s_U can be calculated by Formulas (12) and (14), i.e. $s_L = 54$ and $s_U = 79$, respectively. When $n = 80$, for a given s (s = 1, 2, ..., 80), $G(s)$ can be derived by Formula (11), *i.e.*

$$G(s) = \int_0^{p_0} f(p \mid a+s, b+n-s)dp$$

$$= \int_0^{0.9} B(500+54, 50+80-54)dp$$
$$= 0.9501$$

Similarly,

$$G(55) = 0.9357, \cdots, G(79) = 0.0463,$$
$$G(80) = 0.0340$$

The results of $G(s)$ are shown in Figure 4. Figure 5 shows the sample space. When $n \le 84$, the values of s_L and s_U are given in Table 1.

According to the decision criteria described in section 3.2, when $n = 80$, and $s_L = 54$ $s_U = 79$, the decisions can be made as

- If $s \le 54$, then stop testing and reject H_0.
- If $s \ge 79$, then stop testing and accept H_0.
- If $54 < s < 79$, then continue testing until to a given sample number N or $s \ge s_U$ or $s \le s_L$, and make decisions as described in section 3.3.

As discussed in section 3.2, when $\alpha + \beta < 1$, for a given n, there will not be a theoretical censored sample number N. This is true as discussed in the above case example as shown in Figure 5. However, when n reaches at the proposed censored sample number N, *i.e.* $n = N$, the sequential test must be terminated and a decision has to be made as described in section 3.3. In other words, even if

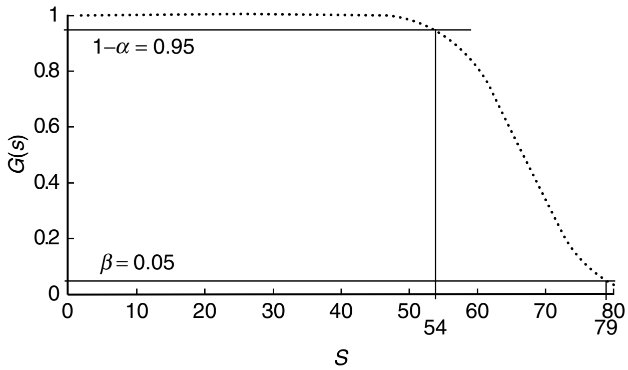

Fig. 4. Results of $G(s)$, when $n = 80$.

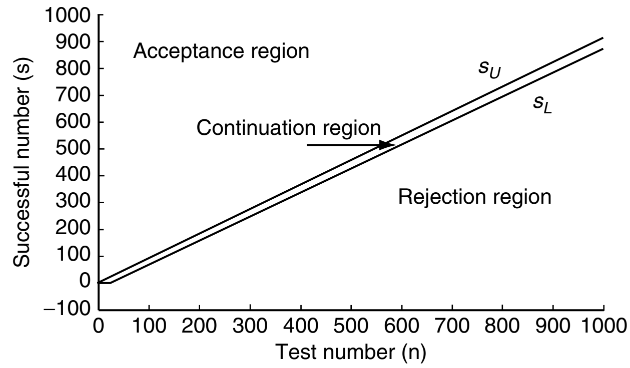

Fig. 5. Sequential tests of binomial distributed products of BSBTM.

Table 1.
Values of S_L and S_U of BSBTM

n	S_L	S_U	n	S_L	S_U	n	S_L	S_U	n	S_L	S_U	n	S_L	S_U
0	–1	1	17	–1	18	34	13	35	51	28	52	68	43	68
1	–1	2	18	–1	19	35	13	36	52	29	53	69	44	69
2	–1	3	19	–1	20	36	14	37	53	29	54	70	45	70
3	–1	4	20	0	21	37	15	38	54	30	55	71	45	71
4	–1	5	21	1	22	38	16	39	55	31	56	72	46	72
5	–1	6	22	2	23	39	17	40	56	32	57	73	47	73
6	–1	7	23	3	24	40	18	41	57	33	58	74	48	74
7	–1	8	24	4	25	41	19	42	58	34	59	75	49	75
8	–1	9	25	5	26	42	20	43	59	35	60	76	50	76
9	–1	10	26	5	27	43	21	44	60	36	61	77	51	77
10	–1	11	27	6	28	44	21	45	61	37	62	78	52	77
11	–1	12	28	7	29	45	22	46	62	37	63	79	53	78
12	–1	13	29	8	30	46	23	47	63	38	64	*80*	*54*	*79*
13	–1	14	30	9	31	47	24	48	64	39	65	81	54	80
14	–1	15	31	10	32	48	25	49	65	40	66	82	55	81
15	–1	16	32	11	33	49	26	50	66	41	66	83	56	82
16	–1	17	33	12	34	50	27	51	67	42	67	84	57	83

the sequential test is carried out continuously, no final result can be received unless a certain censored sample number N is set up to stop testing. Therefore, in order to get a theoretical censored sample number N, a modified BSBTM is needed.

4. Modified Bayesian sequential binomial test model (MBSBTM)

4.1. Fundamentals of MBSBTM

As discussed in section 3.4, a modified BSBTM is needed to solve the problem of no final result in the sequential test to be obtained when $\alpha + \beta < 1$. Considering the 'small probability event principle' (Li and Li, 2008), the following assumptions are applied.

(1) If H_0 is accepted, then for a given (n, s), the posterior probability of $P(p \geq p_0 | (n, s))$ may be near 1, i.e. $P(p \geq p_0 | (n, s)) \approx 1$. In other words, $P(p \geq p_0 | (n, s)) = 1 - P(p \geq p_0 | (n, s)) \approx 0$. If $P(p < p_0 | (n, s))$ is a sufficiently small value, the producer must be sure that the consumer fully accepts the null hypothesis H_0. As described in section 3.1, when $P(p < p_0 | (n, s)) < \beta$, then stop testing and accept H_0. In this case, the probability of Type II error is smaller than β.

Assume $\varepsilon > 0$ is a small value (it should be noted that the aim of introduction of ε is to protect the consumer's benefit). From the viewpoint of consumers, it would be interested in p having a maximum value of probability. Consider $P(p \geq p_0 + \varepsilon | (n, s))$ has a sufficiently large value, e.g. $P(p \geq p_0 + \varepsilon | (n, s)) \approx 1$. Because $P(p \geq p_0 | (n, s)) > P(p \geq p_0 + \varepsilon | (n, s))$, if $P(p \geq p_0 + \varepsilon | (n, s)) \approx 1$, then $P(p \geq p_0 | (n, s)) \approx 1$. In this case, the consumer may be happy to fully accept H_0. For example, when $\alpha = 0.05$, $p_0 = 0.9$, the prior distribution is $B(500, 50)$ and the result of actual and successful sample numbers are (80, 54), the posterior probability of H_0 is $P(p \geq 0.9 | (80, 79)) = 0.9537$. On the basis of decision criteria of BSBTM as described in section 3.1, H_0 cannot be rejected. But by adding a small value of $\varepsilon = 0.025$ into p_0, the posterior probability will be $P(p \geq 0.9 + 0.025 | (80, 79)) = 0.3018$. As can be seen that the posterior probability decreases significantly. Therefore, the following decision criterion is introduced.

(2) If $P(p \geq p_0 + \varepsilon | (n, s))$ is a sufficiently small value, for example, $P(p \geq p_0 + \varepsilon | (n, s)) < 0.05$, then the consumer can reject null hypothesis H_0. As described in section 3.1, in this case, if $P(p \geq p_0 + \varepsilon | (n, s)) < \alpha$, then stop testing and reject H_0. The probability of Type I error is bigger than α because $P(p \geq p_0 | (n, s)) > P(p \geq p_0 + \varepsilon | (n, s))$. ε can be assigned to an agreed value by both consumer and producer, *e.g.* $\varepsilon < 1 - p_0$ based on the particular cases.

On the basis of the above two assumptions, the decision in the sequential test can be naturally as

(3) If $P(p < p_0 | (n, s)) \geq \beta$ and $P(p \geq p_0 + \varepsilon | (n, s)) \geq \alpha$, then continue testing. In this case, the continuation region will be the region in which for any (n, s), the posterior probability of p satisfies $P(p_0 < p < p_0 + \varepsilon | (n, s)) \leq 1 - \alpha - \beta$.

For a given prior distribution $B(a, b)$ and actual sample number (n, s), the variance of p is defined

$$Var(p \mid a+s, b+n-s) = \frac{(a+s)(b+n-s)}{(a+b+n)^2(a+b+n+1)} \quad (15)$$

With n increasing, the variances will be gradually reduced and the PDF of $\pi (p | (n, s))$ will be more and more close. Therefore, a maximum sample number N must exist. When $n = N$, $s = s_0$, $P(p < p_0 | (N, s_0)) \leq \beta$ and $P(p \geq p_0 + \varepsilon | (N, s_0)) \geq \alpha$. Let $s_L = s_U = s_0$, then sample testing must be stopped and the decision of accepting or rejecting H_0 must be made.

4.2. Solution of MBSBTM

When the prior distribution of p is $B(a, b)$, $P(p < p_0 | (n, s)) < \beta$ will be

$$\int_0^{p_0} f\left(p \mid a+s, b+n-s\right) dp < \beta \quad (16)$$

It should be noted that Inequality (16) is same as Inequality (13). Similarly, for any given n, s_U can be calculated by Formula (14).

When $P(p \geq p_0 + \varepsilon | (n, s)) < \alpha$, it becomes

$$1 - \int_0^{p_0+\varepsilon} f\left(p \mid a+s, b+n-s\right) dp < \alpha \quad (17)$$

or

$$\int_0^{p_0+\varepsilon} f\left(p \mid a+s, b+n-s\right) dp \geq 1 - \alpha \quad (18)$$

As can be seen that Inequality (18) is similar to Inequality (10). The only difference is that the upper limit of integral is changed from p_0 to $p_0 + \varepsilon$. The calculation of s_L is similar to BSBTM. But the lower boundary of s_L of MBSBTM may be greater than that of BSBTM (when $\varepsilon > 0$) or equal to that of BSBTM (when $\varepsilon = 0$). Obviously, when $\varepsilon > 0$, the curve of lower boundary s_L will intersect wherever upper boundary s_U at point A as shown in Figure 7. Therefore, the continuation-sampling region will be closed and a sequential test procedure will lead fairly rapid to a terminal decision.

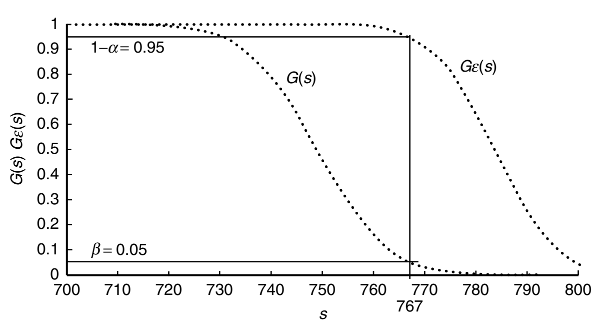

Fig. 6. Results of $G(s)$ and $G_\varepsilon(s)$, when $n = 838$.

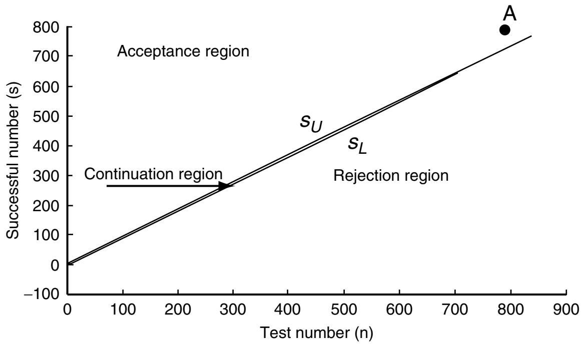

Fig. 7. Sequential test of binomial distributed products of MBSBTM.

In MBSBTM, $G_\varepsilon(s)$ is defined as

$$G_\varepsilon(s) = \int_0^{p_0+\varepsilon} f\left(p \mid a+s, b+n-s\right) dp \qquad (19)$$

which is similar to Formula (12), s_L can be calculated by

$$S_L = \max\{-1, \max\{s \mid G_\varepsilon(s) \geq 1-\alpha\}, s = 1, 2, \cdots, n\} \qquad (20)$$

For $n = 0, 1, 2, \cdots$, s_L and s_U can be calculated by Formulas (20) and (14). The censored sample number N can be obtained

$$N = \min\{n \mid s_L = s_U, n = 0, 1, 2, \cdots n\} \qquad (21)$$

Once s_L, s_U and N are obtained, the decision criteria of sequential tests can be determined.

- If $s_U \leq s \leq n$, then stop testing and accept H_0.
- If $0 \leq s \leq s_L$, then stop testing and reject H_0.
- If $s_L \leq s \leq s_U$, then continue testing until to a maximum sample number N or $s \geq s_U$ or $s \leq s_L$.

4.3. Numerical example 2

Using the data associated with Example 1 as described in Section 3.4, supposing $\varepsilon = 0.025$ and $n = 1, 2, \cdots$, s_U and s_L can be calculated by Formulas (14) and (20). By using Formula (21) a maximum censored sample number N can be determined. Figure 6 shows the results of $G(s)$ and $G_\varepsilon(s)$ when $n = 838$ and $s = 700$ to 838. As can be seen from Figure 6, when $s = 767$, both curves of $G(s)$ and $G_\varepsilon(s)$ satisfy Formulas (14) and (20). Therefore, the censored sample number $N = 838$ when $s_L = s_U = 767$.

Figure 7 shows the sample space. When $n \leq 84$, the values of s_L and s_U are listed in Table 2. Comparing Tables 2 with 1, it can be seen that s_L of some same n in Table 2 are greater than those in Table 1. For example, when $n = 51$, $s_L = 44$ of MBSBTM in Table 2 is greater than $s_L = 28$ of BSBTM in Table 1.

Therefore, the decision criteria are

(1) When $n < 838$
- If $s_U \leq s \leq n$, stop testing and accept H_0.
- If $0 \leq s \leq s_L$, stop testing and reject H_0.
- If $s_L < s < s_U$, continue testing.

(2) When $n = 838$
- If $s > 767$, stop testing and accept H_0.
- If $s \leq 767$, stop testing and reject H_0.

Table 2.
Numerical values of S_L and S_U of MBSBTM

n	S_L	S_U	n	S_L	S_U	n	S_L	S_U	n	S_L	S_U	n	S_L	S_U
0	–1	1	17	13	18	34	29	35	51	44	52	68	60	68
1	–1	2	18	14	19	35	30	36	52	45	53	69	61	69
2	–1	3	19	15	20	36	30	37	53	46	54	70	62	70
3	0	4	20	16	21	37	31	38	54	47	55	71	62	71
4	1	5	21	17	22	38	32	39	55	48	56	72	63	72
5	2	6	22	18	23	39	33	40	56	49	57	73	64	73
6	3	7	23	19	24	40	34	41	57	50	58	74	65	74
7	4	8	24	19	25	41	35	42	58	51	59	75	66	75
8	5	9	25	20	26	42	36	43	59	51	60	76	67	76
9	6	10	26	21	27	43	37	44	60	52	61	77	68	77
10	7	11	27	22	28	44	38	45	61	53	62	78	69	77
11	8	12	28	23	29	45	39	46	62	54	63	79	70	78
12	8	13	29	24	30	46	40	47	63	55	64	80	71	79
13	9	14	30	25	31	47	40	48	64	56	65	81	72	80
14	10	15	31	26	32	48	41	49	65	57	66	82	73	81
15	11	16	32	27	33	49	42	50	66	58	66	83	73	82
16	12	17	33	28	34	50	43	51	67	59	67	84	74	83

It should be noted that the prior distribution of p should be feasible and accurate. The more accurate the prior distribution is, the more correction the deduction of p will be. Therefore, the prior information may need to be collected as much as possible. These include historical data analysis, failure analysis, concept mapping, and domain human expert experience and engineering knowledge analysis. However, if the rejection or acceptance conclusions are obtained by adopting the prior distribution, the decision needs to be made whether the actual sequential test is needed or not. For example, if $p_0 = 0.90$, $\alpha = \beta = 0.05$ and prior distribution is B (645,55), then

$$\int_0^{p_0} f\left(p\,|\,645{,}55\right) dp = 0.0231 < 0.05$$

On the basis of the Assumption 1) of BSBTM as described in section 3.1, the decision of accepting H_0 can be made, and no actual test is needed. However, if the prior distribution is B(615, 55), then

$$\int_0^{p_0} f\left(p\,|\,615{,}85\right) dp = 0.9644 > 1 - 0.05$$

In this case, the decision of rejecting H_0 can be made based on the Assumption 2) of BSBMT as described in section 3.1, and no actual test is needed.

4.4. Simulation of average sample number (ASN)

The MBSBTM introduces a small positive value ε to be assigned, the maximum censored sample number N can be obtained by Formula (21). In fact, N will be the maximum sample numbers for the binomial sequential test.

As discussed earlier in this paper, in practice, a Bayesian sequential test could be stopped at a time where s satisfies the decision criteria. Therefore, the actual sample number n is often smaller than or equal to the maximum sample number N in hand, i.e. $n \leq N$. In order to make an effective and efficient testing plan, it would be useful if the expected sample number of a sequential test, i.e. the average sample number (ASN), can be predicted in advance.

The ASN can be obtained by adopting a simulation method which involves the following steps:

Step 1: For given $B(a, b)$, ε and N, calculate s_L and s_U ($n = 0, 1, 2, \cdots, N$). The sample space can be obtained as shown in Figure 7.

Step 2: Establish simulation criterion by assigning the absolute precision ε_A, for example, $\varepsilon_A = 0.0001$.

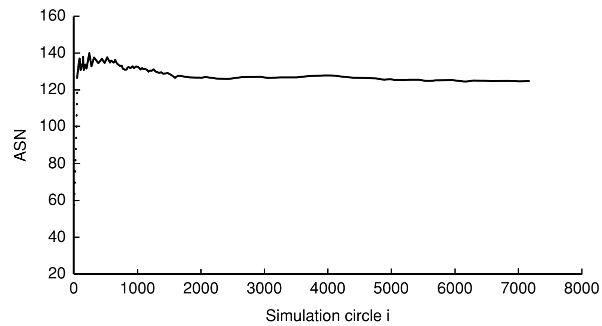

Fig. 8. Simulation process of ASN calculation.

Step 3: Let $i: = 1$, where i denotes the i^{th} binomial sequential test.

Step 4: Let $sum: = 0$, where sum denotes the total sample number of the previous i^{th} test.

Step 5: Let $ASN_0: = 0$.

Step 6: The i^{th} simulation process

- *Step 6.1*: Let $n_i : = 0$ and $s_i : = 0$, where n_i is the actual successful sample number in the i^{th} test and $n_i \leq N$.
- *Step 6.2*: If $s_i \geq s_U$ and $s_i \leq s_L$, then stop simulation and go to Step 7, otherwise, let $n_i : = n_i + 1$.
- *Step 6.3*: Calculate the random number p_{ni} based on the prior distribution $B(a,b)$.
- *Step 6.4*: Let p_{n_i} be the probability of success ratio of the $n_i{}^{\text{th}}$ test.
- *Step 6.5*: Calculate the random number t based on the binomial distribution $b(1, p_{ni})$ and let $s_i: = s_i + t$, where t denotes the result of the $n_i{}^{\text{th}}$ test ($t = 0$ - test failure and $t = 1$ - test success).
- *Sept 6.6*: Go to Step 6.2.

Step 7: Let $sum: = sum + n_i$.

Step 8: Calculate the ASN_i.

Step 9: If $|ASN_i - ASN_{i-1}| < \varepsilon_A$, then stop simulation process, let $ASN: = ASN_i$ and go to Step 11.

Step 10: If $|ASN_i - ASN_{i-1}| \geq \varepsilon_A$, then let $i: = i + 1$ and go to Step 6.

Step 11: End

A Matlab programme for the simulation has been developed following the steps as described above. Using the data associated with case Examples 1 and 2 as described in Sections 3.4 and 4.3, the ASN are calculated. In this case, $\varepsilon_A = 0.0001$. When i = 11886, the simulation is terminated and ASN = 126.0612. Figure 8 shows the simulation process.

5. Conclusion

Testing the lifetimes of components or products by means of the verification of probability of success ratio on which to base a statistical probability of the binomial distribution is often a costly and difficult undertaking. The current methods such as SPRT and SPOT are comparatively mature methods and usually depend on a hypothesis testing of p resulting in poor performance with increasing costs and time delay in the sequential reliability tests, even cannot reach at a desirable target. Therefore, it is essential to develop new methods to predict the sample numbers and to verify of success ratio p in an acceptable way under various environments where such mature methods cannot be effectively or efficiently applied. This paper presents the BSBTM and MBSBTM methodologies for verification of probability of success ratio and prediction of ASN for binomial distributed components or products in the sequential tests. This paper also presents two case examples to demonstrate the proposed methods. Comparing with those traditional methods, the advantages of the proposed methods can be summarized (1) for a given prior distribution, $B(a, b)$, the proposed methods take the posterior probability (or α and β) into consideration, which enables a good testing plan can be made before the implementation of tests. By using such a testing plan, the testing process can be monitored and sequential test decision can be made effectively and efficiently, (2) by using MBSBTM, the average simple number (ASN) can be predicted in advance before the test implementation, (3) the proposed BSBTM and MBSBMT are more convenient to incorporate into a computer programmes such as Matlab, and (4) the introduction of ε based on the 'small probability event principle' enables decision being made to perform a sequential test of the underlying actual testing environment so that more reliable result can be obtained.

Acknowledgement

The work described herein is part of a research project on reliability analysis funded by National Science Foundation of China (Grant No. 70971133). Their support is gratefully acknowledged.

References

Barnett V. D., 1972. Bayesian sequential life test. *Technometrics* **14(2)**, 453–468.

Dekking F.M., Kraaikamp C., Lopuhaa H.P. and Meester L.E., 2005. A modern introduction to probability and statistics: Understanding why and how. Springer, London.

Desu M.M. and Raghavarao D., 1990. Sample size methodology. Academic Press, Boston.

Ghosh B.K., 1970. Sequential tests of statistical hypotheses. Reading (Mass). Addison-Wesley, London.

Ghosh J.K., Delampady M. and Samanta T., 2006. An introduction to Bayesian analysis: Theory and methods. Springer, New York.

Guo B., Jing P. and Xing Y.Y., 2008. A censored sequential posterior odd test (SPOT) method for verification of the mean time to repair. *Reliability, IEEE Transactions* **57(2)**, 243–247.

Hamada M.S., Wilson A.G., Reese C.S. and Martz H.F., 2008. Bayesian reliability. Springer Science and Business Media, New York.

Hines W.W. and Montgomery D.C., 2003. Probability and statistics in engineering and management science. Hoboken, NJ: Wiley, 4th ed.

Hoff P.D., 2009. A first course in Bayesian statistical methods. Springer, New York.

Jiang G., Li C. and Liu B., 2009. Test design of operational applicability for torpedo based on Bayes SPRT method. *Torpedo Technology* **17**, 77–80.

Kapur K.C. and Lamberson L.R., 1977. Reliability in engineering design. John Wiley and Sons, Inc., New York.

Kin C.C., Seraidaridia C.I. and Tortorella M., 1997. Sequential sampling plans for early-life reliability assessment. Proceedings of Annual Reliability and Maintainability Symposium, 131–135, London.

Koch K.R., 2007. Introduction to Bayesian statistics., Springer-Verlag Berlin Heidelberg, Berlin, 2nd ed.

Lefebvre M., 2006. Applied probability and statistics. Springer, New York.

Li X.R. and Li X.B., 2008. Common fallacies in applying hypothesis testing. Proceedings of the 11th International Conference on Information Fusion, 1–8, New York.

Wetherill G.B. and Glazebrook K.D., 1986. Sequential methods in statistics. Chapman and Hall, London, 3rd ed.

ACRONYM

BSBTM	Bayesian sequential binomial test model
MBSBTM	Modified Bayesian sequential binomial test model
ASN	Average sample number

SPRT	Sequential probability ratio test
SPOT	Sequential posterior odd test
PDF	Probability density function

NOTATION

$B(a, b)$	Beta distribution (a and b are known)
α	Risk of producer (probability of Type I error)
β	Risk of consumer (probability of Type II error)
n	Sample size, *i.e.* current actual test number or field test number
s	Current successful sample number
s_L	Lower boundary of continuation sampling region
s_U	Upper boundary of continuation sampling region
N	Maximum actual sample number
p	Success ratio
p_0	Given system reliability requirement
H_0	Null hypothesis in hypothesis test, *i.e.* hypothesis of $p \geq p_0$
H_1	Alternative hypothesis in hypothesis test, *i.e.* hypothesis of $p < p_0$
ε	A small positive value
$f(p \mid a, b)$	PDF of p, and a, b are known parameters
$P(H_0 \mid x)$	Posterior probability of H_0, *i.e.* $P(p \geq p_0 \mid x)$, where x is the actual test result
$P(H_1 \mid x)$	Posterior probability of H_1, *i.e.* $P(p < p_0 \mid x)$

Appendix

$$G(s) = P(H_1 \mid (n,s)) = P(p < p_0 \mid (n,s)) \tag{22}$$

For given a, b and n, $G(s)$ is a strictly monotone decreasing function of s, where $s = 0, 1, 2, ..., n$.
Proof
Let $g(s) = p^{s+1} (1-p)^{n-s-1}$, according to Bayesian Formula (5)

$$G(s) = \frac{\int_0^{p_0} \frac{1-p}{p}\pi(p)g(s)dp}{\int_0^1 \frac{1-p}{p}\pi(p)g(s)dp}, \quad G(s+1) = \frac{\int_0^{p_0} \pi(p)g(s)dp}{\int_0^1 \pi(p)g(s)dp}$$

then

$$\begin{aligned}
\frac{G(s)}{G(s+1)} &= \frac{\int_0^{p_0} \frac{1-p}{p}\pi(p)g(s)dp \int_0^1 \pi(p)g(s)dp}{\int_0^1 \frac{1-p}{p}\pi(p)g(s)dp \int_0^{p_0} \pi(p)g(s)dp} \\
&= \frac{\int_0^{p_0} \frac{1-p}{p}\pi(p)g(s)dp \left(\int_0^{p_0} \pi(p)g(s)dp + \int_{p_0}^1 \pi(p)g(s)dp \right)}{\left(\int_0^{p_0} \frac{1-p}{p}\pi(p)g(s)dp + \int_{p_0}^1 \frac{1-p}{p}\pi(p)g(s)dp \right) \int_0^{p_0} \pi(p)g(s)dp} \\
&= \frac{\int_0^{p_0} \frac{1-p}{p}\pi(p)g(s)dp \int_0^{p_0} \pi(p)g(s)dp + \int_0^{p_0} \frac{1-p}{p}\pi(p)g(s)dp \int_{p_0}^1 \pi(p)g(s)dp}{\int_0^{p_0} \frac{1-p}{p}\pi(p)g(s)dp \int_0^{p_0} \pi(p)g(s)dp + \int_{p_0}^1 \frac{1-p}{p}\pi(p)g(s)dp \int_0^{p_0} \pi(p)g(s)dp}
\end{aligned}$$

Because $\frac{1-p}{p} = \frac{1}{p} - 1$ is a strictly monotone decreasing function of p, when $0 < p_0 < 1$

$$\int_0^{p_0} \frac{1-p}{p} \pi(p) g(s) dp \int_{p_0}^1 \pi(p) g(s) dp > \frac{1-p_0}{p_0} \int_0^{p_0} \pi(p) g(s) dp \int_{p_0}^1 \pi(p) g(s) dp$$

$$\int_{p_0}^1 \frac{1-p}{p} \pi(p) g(s) dp \int_0^{p_0} \pi(p) g(s) dp < \frac{1-p_0}{p_0} \int_{p_0}^1 \pi(p) g(s) dp \int_0^{p_0} \pi(p) g(s) dp$$

Therefore

$$\frac{G(s)}{G(s+1)} > \frac{\int_0^{p_0} \frac{1-p}{p} \pi(p) g(s) dp \ \int_0^{p_0} \pi(p) g(s) dp + \frac{1-p_0}{p_0} \int_0^{p_0} \pi(p) g(s) dp \int_{p_0}^1 \pi(p) g(s) dp}{\int_0^{p_0} \frac{1-p}{p} \pi(p) g(s) dp \int_0^{p_0} \pi(p) g(s) dp + \frac{1-p_0}{p_0} \int_{p_0}^1 \pi(p) g(s) dp \int_0^{p_0} \pi(p) g(s) dp} = 1$$

For any $s(s = 0, 1, 2, ..., n)$, $G(s) > G(s + 1)$, $G(s)$ is therefore a strictly monotone decreasing function of s.